IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF TEXAS

HUMANA INC. *and* AMERICANS FOR BENEFICIARY CHOICE,

Plaintiffs,

v.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES; CENTERS FOR MEDICARE & MEDICAID SERVICES; XAVIER BECERRA, *in his official capacity as Secretary of Health and Human Services; and* CHIQUITA BROOKS-LASURE, *in her official capacity as Administrator of the Centers for Medicare and Medicaid Services*,

Defendants.

Case No.: 24-cv-01004-O

COMPLAINT

Plaintiffs Humana Inc. and Americans for Beneficiary Choice, for their complaint under the Administrative Procedure Act (5 U.S.C. §§ 702, 704) against defendants U.S. Department of Health and Human Services; Centers for Medicare & Medicaid Services; Xavier Becerra, in his official capacity; and Chiquita Brooks-LaSure, in her official capacity, allege as follows.

TABLE OF CONTENTS

INTRODUCTION

1. This case is a challenge to the federal government's arbitrary and capricious actions in administering the Medicare Advantage and Part D Star Ratings program. The data and calculations underlying the annual Star Ratings are dizzyingly complex, and at a first glance, this suit may appear to be a dry disagreement over technical details. It is anything but. In fact, it is about enforcing the settled ground rules for agency decisionmaking under the Administrative Procedure Act (APA)—in particular, the Center for Medicare and Medicaid Services' (CMS) duties to follow its own regulations, to put those regulations into practice using reason and logic, and to follow commonsense protocols concerning data integrity and agency transparency.

2. Medicare is a federal health insurance program for seniors and people with disabilities. Enrollees in the program can choose coverage under either "traditional" Medicare or the semi-privatized Medicare Advantage (MA) program. CMS is the federal agency responsible for administering the Medicare program. In that capacity, CMS calculates and publishes a Star Rating (on a scale of one to five stars) for each health benefit plan offered under Medicare Advantage and its companion drug benefit program, Medicare Part D. The Rating is intended to reflect plan quality and performance based on a range of underlying quality measures.

3. Star Ratings are tremendously important to the operation of the MA and Part D programs. First, they provide agents and brokers, and the Medicare beneficiaries they serve, with information about a plan's quality, enabling them to compare plans when shopping during the annual enrollment period. In addition, CMS must provide quality bonus payments—in amounts that can reach hundreds of millions or even billions of dollars annually—to plans with better Star Ratings. Plans must then use those payments either to lower costs for their enrollees or to provide them with additional benefits.

4. The stakes hardly could be higher. Last year and for the first time, Medicare Advantage surpassed traditional Medicare measured by its share of the 60+ million Americans who depend on Medicare. It is now a *half-trillion-dollar* public benefit program.

5. Plaintiff Humana Inc. is one of the nation's largest MA organizations (MAOs), or sponsors of health insurance plans under MA and Part D. It is committed to putting health first by designing and administering MA plans of the highest quality. Indeed, high-quality healthcare and high-quality service have been the primary drivers of Humana's success over its three decades participating in the Medicare programs for private health plans, leading to a better quality of life for the enrollees it serves. The high quality of the plans sponsored by Humana is reflected in the industry-leading Star Ratings they have been assigned over the past six years. And precisely because Humana is committed to quality, it also is committed to the integrity of the Star Ratings system.

6. Plaintiff Americans for Beneficiary Choice (ABC) is a non-profit trade association whose members include the agents and brokers who use the Star Ratings system to make informed recommendations, as well as the beneficiaries to whom they sell MA and Part D plans. All of the participants in the complex Medicare Advantage industry, including ABC's members, are guided by Star Ratings and depend on them to be reliable and accurate.

7. On October 10, 2024, CMS finalized and released the 2025 Star Ratings. Across the board, the number of MA plans with high Star Ratings decreased significantly year-over-year. Under the 2023 Star Ratings, 21.87% of MA participants had been enrolled in 5.0 Star plans. In the 2024 Ratings, that number decreased markedly, to 7.64%. And this year, for the 2025 Star Ratings, the number plummeted yet further, to a vanishingly small 1.79%. Meanwhile, the number of enrollees in 3.5 Star plans ballooned from 18.71% in 2023 and 15.89% in 2024, to 27.71% in 2025. There are no broader, objective indications that MA plan quality has diminished over that time period.

8. In the runup to the release of the 2025 Star Ratings, Humana actively participated in CMS's "plan preview periods," during which the agency requires MA organizations to help verify the Star Ratings data and calculations. What Humana uncovered was concerning. The measure-level "cut points" (the four thresholds used to convert raw numerical scores into measure-level Star Ratings) historically have remained steady year-on-year. Indeed, recent changes to the agency's methodology were designed to promote cut-point stability, by removing outliers from the agency's data. But this year, the cut points for several measures moved abruptly and substantially upward, significantly depressing MAOs' Star Ratings, including Humana's. Under the 2024 Ratings, 94% of Humana's MA enrollees were in a plan with 4.0 Stars or higher. As a result of the unexplained swings in the most recent cut points calculated by CMS, now only 25% of its enrollees are in plans rated 4.0 stars and above for 2025.

9. Humana was denied an opportunity to determine why the measure-level cut points moved so suspiciously in the 2025 scores or to validate the accuracy of CMS's calculations. Despite that CMS's own regulations call for disclosure to MA plan sponsors of the data underlying CMS's Star Ratings calculations, the agency refused to share information necessary for Humana to verify the agency's work in time to make corrections.

10. In addition, the Star Ratings calculations include measures to evaluate the performance of plans' customer-service call centers in providing foreign-language interpreters to assist would-be enrollees calling to seek information about plan benefits. To assess these measures, CMS conducts the Accuracy & Accessibility Study, where call "surveyors" place test calls to evaluate centers' compliance with regulatory requirements. Here, CMS lowered the Star Ratings for at least a dozen of Humana's largest plans on the basis of just three phone calls that were handled by CMS in a manner inconsistent with the agency's own regulations.

11. Humana, ABC, and ABC's members all count on CMS to administer the Star Ratings system in a consistent, transparent, and rational manner that accords with its own regulations. If the agency develops its methodologies and undertakes its calculations in a black box while refusing to allow MA organizations and other third parties to validate its work, neither regulated plans nor Medicare beneficiaries and their third-party agents and brokers will be able to rely with any confidence on the agency's reported results—least of all in years like this, when the agency's calculations take substantial and unexplained swings that align more readily with the agency's interest in reducing payments than actual plan quality.

12. The APA requires federal agencies follow their own rules, be open with their data and reasoning, and provide logical explanations for their decisions. CMS did none of that with respect to the 2025 Star Ratings, Humana's included. Humana and ABC thus bring this action seeking an order vacating Humana's 2025 Star Ratings and remanding the matter to the agency to recalculate the Ratings in accordance with its own regulations and with the transparency necessary for all MA organizations to assist the agency—and to ensure the integrity of the system—by validating its work.

13. In light of the typical schedule for new contract and service-area expansion applications, as well as the time that is needed to prepare annual bids, the claims presented here are matters of pressing concern, warranting expedited resolution by the Court. The 2025 Star Ratings will play an important role in every MAO's calculation of its bids for contract year (CY) 2026. As a result, Humana (and all MAOs) must have access to the data that CMS so far has refused to disclose by mid-December 2024. Only then will they have adequate time to analyze the data so that it can make informed bids for CY 2026.

14. It is possible to retract and recalculate unlawful Star Ratings even mid-contract-year, as CMS did after another court held in expedited proceedings that the agency

had committed a systemic legal error in the 2024 Star Ratings. *See SCAN Health Plan v. HHS*, 2024 WL 2815789 (D.D.C. June 3, 2024). The same expedited relief and outcome is warranted here.

JURISDICTION AND VENUE

15. This action arises under the APA, 5 U.S.C. §§ 702, 704; and the Declaratory Judgment Act, 28 U.S.C. § 2201. The Court's subject-matter jurisdiction is invoked under 28 U.S.C. § 1331.

16. Venue is proper in this District under 28 U.S.C. § 1391(e) because this is an action against officers and agencies of the United States not involving real property, and one of the plaintiffs resides in this District.

PARTIES

17. Plaintiff Humana Inc. is a Delaware corporation with its principal place of business at 500 West Main Street, Louisville, Kentucky 40202. Humana and its subsidiaries are providers of healthcare services, with approximately 17 million health-plan participants across the United States.

18. Plaintiff Americans for Beneficiary Choice (ABC) is a trade association based in Dallas, Texas. ABC's members include health insurance industry leaders and workers, consumer advocates, and concerned citizens. ABC's mission is to protect the best interests of Medicare and other health insurance beneficiaries through legislative and regulatory advocacy and participation in litigation. Through these efforts, it aims to improve the American healthcare system with sensible, forward-thinking policies that improve health insurance knowledge and education, lower healthcare costs, and maximize coverage choice for consumers. The interests and objectives that ABC seeks to advance in this litigation are thus directly relevant to its institutional mission.

19. The U.S. Department of Health and Human Services (HHS) is a cabinet-level agency within the United States government. Xavier Becerra, sued in his official capacity, is the Secretary of HHS. Congress has assigned HHS ultimate responsibility for administering the Medicare Advantage and Medicare Part D programs.

20. HHS has delegated authority to administer the Medicare Advantage and Medicare Part D programs to the Centers for Medicare & Medicaid Services (CMS), an agency within HHS. *See* 66 Fed. Reg. 35437. Defendant Chiquita Brooks-LaSure, sued in her official capacity, is Administrator of CMS. CMS manages the Star Ratings system and issued the Star Ratings decision that is the final agency action challenged in this case.

STATUTORY AND REGULATORY BACKGROUND

A. The Medicare Advantage and Medicare Part D programs

21. Established in 1965 as an amendment to the Social Security Act, the federal Medicare program is the federal health insurance program for people aged 65 or older or with certain disabilities or end-stage renal disease. *See Medicare Program; Establishment of the Medicare Advantage Program*, 70 Fed. Reg. 4588 (Jan. 28, 2005); 42 U.S.C. § 1395kk(a).

22. Medicare comprises four parts: Parts A, B, C, and D. *See* 70 Fed. Reg. at 4589. Medicare Part A (which covers inpatient hospital treatment) and Part B (which covers outpatient services) are together known as "traditional" or "original" Medicare. Traditional Medicare use a fee-for-service payment model. *See* 42 U.S.C. § 1395w-22(a)(1). CMS thus reimburses providers directly for the services they provide to traditional Medicare beneficiaries. *MaxMed Healthcare, Inc. v. Price*, 860 F.3d 335, 337 (5th Cir. 2017); *UnitedHealthcare Insurance v. Becerra*, 16 F.4th 867, 872 (D.C. Cir. 2021).

23. Medicare Part C, also known as Medicare Advantage, uses a different model. *See* Medicare Prescription Drug, Improvement, and Modernization Act, Pub. L. No. 108-

173, 117 Stat. 2066 (2003) (codified at 42 U.S.C. §§ 1395w-21 to 1395w-28). The program avoids the pitfalls of traditional Medicare and its single-payer, one-size-fits-all approach by offering plans sponsored by private companies called Medicare Advantage organizations, or MAOs. These companies must cover at least the same services that Medicare beneficiaries would receive through traditional Medicare. 42 U.S.C. § 1395w-22(a). But to attract enrollees, MA plans typically offer additional benefits not covered by traditional Medicare, such as dental and vision insurance. *UnitedHealthcare*, 16 F.4th at 872.

24. Under this public-private partnership model, MAOs do not receive fee-for-service reimbursements from CMS for the healthcare services their enrollees receive. *See generally* 42 U.S.C. § 1395w-23(a). Instead, they receive a per-enrollee monthly payment to provide coverage for all Medicare-covered benefits to the beneficiaries enrolled in their plan. *Id.* In turn, MAOs pay healthcare providers for the services they provide to MA enrollees. *Id.* § 1395w-23(a)(1); *see Caris MPI v. UnitedHealthcare, Inc.*, 108 F.4th 340, 344 (5th Cir. 2024).

25. CMS determines a plan's monthly payment by comparing the plan's "bid" (its estimated cost of providing Medicare-covered services to a particular patient population) to a "benchmark" (the maximum amount the federal government will pay to provide coverage in the plan's service area). *Id.* § 1395w-23(b)(1)(B), (n).

26. If the MAO's bid is below the benchmark, CMS pays the MAO its bid rate, while also returning a specified percentage of the difference between the benchmark and the bid as a "rebate," which must be used to provide additional benefits or otherwise returned to plan participants through lower premiums or cost sharing. 42 U.S.C. §§ 1395w-23(a)(1)(B)(i), (E); 1395w-24(b)(1)(C).

27. If, in contrast, an MAO's plan bid is at or above the benchmark, the MAO receives monthly payments at the benchmark rate, and the MAO must charge enrollees an

additional premium to cover the amount by which the bid exceeds the benchmark. *Id.* §§ 1395w-23(a)(1)(B)(ii), 1395w-24(b)(2)(A). *See also Medicaid & Medicare Advantage Products Association of Puerto Rico v. Emanuelli Hernández*, 58 F.4th 5, 8 n.1 (1st Cir. 2023); *Elevance Health inc. v. Becerra*, 2024 WL 2880415, at *2 (D.D.C. June 7, 2024).

28. In addition to inpatient treatment and outpatient services, Medicare beneficiaries may also obtain prescription drug coverage through Medicare Part D. Like Medicare Advantage, the Part D prescription drug benefit provides coverage through a public-private partnership with plan sponsors. These plan sponsors offer both standalone prescription drug plans (PDPs) for individuals enrolled in traditional Medicare and drug coverage bundled with an MA plan, known as an MA-PD plan. 42 U.S.C. § 1395w–101(a)(1), (3)(C).

29. The enriched range of consumer options introduced by the MA program has produced commensurate decision-making complexity for Medicare beneficiaries who are considering enrolling in an MA plan. Congress intended for insurance brokers and agents to assist Medicare beneficiaries with their decisionmaking in this space. *See* 42 U.S.C. § 1395w-21(j)(2)(D). Indeed, agents and brokers help "millions of Medicare beneficiaries to learn about and enroll in" MA plans "by providing expert guidance on plan options in their local area, while assisting with everything from comparing costs and coverage to applying for financial assistance." 89 Fed. Reg. at 30617.

30. Under the model that is prevalent across the MA program, agents and brokers are unaffiliated with, and not beholden to, MAOs. As independent agents, they can offer beneficiaries a diverse array of MA plans to best meet beneficiaries' needs.

31. Since its adoption by the Bush administration in 2003, the MA and Part D programs have grown steadily. Americans prefer the choices that Medicare Advantage plans provide compared with traditional Medicare. The immediate predecessor to MA, called Medicare + Choice, had approximately 1.56 million enrollees in 1992. *See* CMS,

Medicare Managed Care Contract (MMCC) Plans Monthly Report, https://perma.cc/YPK6-DDEW (click Live View). By 2023, that figure had increased to more than 30 million enrollees, surpassing for the first time the number of beneficiaries opting for traditional Medicare. Nancy Ochieng, et al., *Medicare Advantage in 2023: Enrollment Update and Key Trends* (Aug. 9, 2023), https://perma.cc/EYE2-4UHR. And the Congressional Budget Office recently projected that 62% of Medicare beneficiaries would be enrolled in Medicare Advantage by 2033. Ochieng N. et al., *Medicare Advantage in 2023: Enrollment Update and Key Trends*, Kaiser Family Foundation (Aug. 9, 2023), https://perma.cc/FDQ5-8C36.

32. Recognizing the importance of public participation in the rules and policies governing the MA and Part D programs, Congress specified by statute that "[n]o rule, requirement, or other statement of policy (other than a national coverage determination) that establishes or changes a substantive legal standard governing the scope of benefits, the payment for services, or the eligibility of individuals, entities, or organizations to furnish or receive services or benefits under this subchapter shall take effect unless it is promulgated by the Secretary by regulation" using notice and comment. 42 U.S.C. § 1395hh(a). This includes rules governing the Star Ratings methodologies.

B. The Star Ratings system and score calculations

33. To assist both agents and brokers and inform would-be enrollees, CMS established the Quality Star Ratings system early in the program's existence. Star Ratings measure the quality of health and drug services received by plan participants enrolled in MA and Part D. *See* 42 C.F.R. §§ 422.162(b)(1), 423.182(b)(1).

34. CMS evaluates MA and Part D plans along a range of quality, compliance, and other measures, and develops ratings on a five-star scale based on these measures. *See id.* §§ 422.166(a)(4), 423.186(a)(4). A 1.0 Star Rating is the worst rating, and 5.0 Star Rating is the best. *Id.* §§ 422.166(a)(4), (c)(3), (d)(2)(iv), 423.186(a)(4), (c)(3),(d)(2)(iv). The

system is intended to reflect the quality and performance of each plan. 42 C.F.R. §§ 422.162(b)(1), 423.182(b)(1); *see also Elevance*, 2024 WL 2880415, at *2.

35. The Star Ratings are based on the scores that these plans earn on various quality and performance "measure[s]." *See* 42 C.F.R. §§ 422.162(a), 423.182(a). CMS looks at measures within five broad categories: (1) outcome measures, which reflect improvements in a beneficiary's health; (2) intermediate outcome measures, which reflect actions taken which can assist in improving a beneficiary's health status; (3) patient experience measures, which reflect beneficiaries' perspectives of the care they received; (4) access measures, which reflect whether the plan creates barriers to beneficiaries receiving needed care; and (5) process measures, which capture the health care services provided to beneficiaries that can assist them in maintaining, monitoring, or improving their health status. *See* CMS, *Medicare 2024 Part C & D Star Ratings Technical Notes* 9, https://perma.cc/-Y7VK-BXN9; *Contract Year 2019 Changes to the Medicare Advantage and Medicare Prescription Drug Benefit Programs*, 83 Fed. Reg. 16440, 16532 (Apr. 16, 2018).

36. Each plan receives a numerical score on its applicable measures, which CMS converts into a "measure-level" Star Rating on a five-star scale using four thresholds or "cut points" to divide the distribution of measure scores into five "whole star increments." 42 C.F.R. §§ 422.166(a)(4), 423.186(a)(4). Because the development of measure-level Star Ratings involves the conversion of fine-grained numerical measure scores into just five rating levels based on where the cut points lie, very small movements in the cut points can lead to dramatic changes in a plan's measure-level Star Ratings, despite stability in the plan's underlying quality and performance. Thus, even small errors in CMS's determination of the cut points can have a profound impact on the measure-level Star Ratings that plans are assigned.

37. From the measure-level Star Ratings, CMS calculates Part C and Part D "summary" ratings, which reflect the weighted mean of a plan's measure-level Star Ratings. *Id.* §§ 422.166(c), 423.186(c). CMS further calculates an overall rating for each MA-PD contract, which reflects the weighted mean of that contract's Part C and Part D measure-level Star Ratings.

38. By statute, the Star Ratings that CMS assigns to an MA plan or Part D plan must be based on the data collected in connection to the "ongoing quality improvement program[s]" that each MAO is required to establish. 42 U.S.C. §§ 1395w-22(e)(1), (3); 1395w-23(o)(4)(A); 1395w-151(b). These data sources include quality-of-care performance measures, which Medicare managed care organizations are required to report annually through the Healthcare Effectiveness Data and Information Set (HEDIS) scheme; measures of beneficiaries' experiences with their health plans drawn from the Consumer Assessment of Healthcare Providers & Systems (CAHPS) survey; and measures of changes in the physical and mental health of MA enrollees captured through the Health Outcomes Survey. *See* 83 Fed. Reg. at 16531. In addition to measures from these data sources, MA plan Star Ratings are also based on performance measures that "address telephone customer service, members' complaints, disenrollment rates, and appeals." *Id.*

C. The purpose and effect of the Star Ratings system

39. The Star Ratings systems serves three purposes, each of which requires the ratings to "accurately . . . reflect true performance." 83 Fed. Reg. at 16519.

40. First, the system is designed to provide Medicare beneficiaries with "comparative information on plan quality and performance," allowing them to make "knowledgeable enrollment and coverage decisions in the Medicare program." 42 C.F.R. §§ 422.160-(b)(1), 423.180(b)(1). As CMS has explained, the "MA and Part D Star Ratings system is designed to provide information to the beneficiary that is a true reflection of the plan's

quality and encompasses multiple dimensions of high-quality care," with the goal of "inform[ing] plan choice" by beneficiaries. 83 Fed. Reg. at 16520. To this end, CMS maintains the Medicare Plan Finder website, which displays information about available plans, including each plan's Star Rating. *Elevance*, 2024 WL 2880415, at *2.

41. Second, the system is designed to help CMS perform "oversight, evaluation, and monitoring of MA and Part D plans" (83 Fed. Reg. at 16520-16521) and compliance with regulatory and contract requirements. 42 C.F.R. §§ 422.160(b)(3), 423.180(b)(3).

42. These two goals were the initial impetus for the Star Ratings system. *See Proposed Changes to the Medicare Advantage and the Medicare Prescription Drug Benefit Programs for Contract Year 2012*, 75 Fed. Reg. 71190, 71219 (Nov. 2, 2010) (specifying the purposes "historically" served by the Star Ratings program as helping beneficiaries make an informed choice when selecting a plan and assisting the agency in identifying poor performance to target for compliance actions).

43. The Star Ratings program's third, more recent, purpose is to provide "quality ratings on a 5-star rating system" to be used in administering the scheme of additional payments for high quality MA plans, known as quality bonus payments (QBPs). The QBP system was established in 2010 by the Patient Protection and Affordable Care Act (ACA). *See* 42 C.F.R. § 422.160(b)(2); 75 Fed. Reg. at 71218.

44. The ACA provides that an MA plan is entitled to QBPs from CMS depending on the "quality rating" of the plan, which "shall be determined according to a 5-star rating system." 42 U.S.C. § 1395w-23(e)(4)(A). Thus, if an MA plan receives a Star Rating of 4 stars or higher, its benchmark amount is increased, in turn increasing the rebates that CMS will pay by increasing the difference between the plan sponsor's benchmark and its bid. *Id.* § 1395w-23(o)(1), (3)(A).

45. Star Ratings also determine the portion of the difference that is returned as a rebate. 42 C.F.R. §§ 422.162(b)(2), 423.182(b)(2). Plans with a 4.5 Star Rating or higher receive 70% of the difference between the benchmark and the bid; plans with a Rating between 3.5 and 4.5 Stars receive a 65% rebate, and plans with a rating under 3.5 stars receive a 50% rebate. 42 U.S.C. § 1395w-24(b)(1)(C)(v); 42 C.F.R. § 422.266(a)(2)(ii).

46. CMS prominently displays Star Ratings in its online and print resources concerning available MA plans. *See* 42 U.S.C. § 1395w-21. Through the online Medicare Plan Finder tool, CMS displays MA plans to prospective enrollees in order of highest to lowest Star Ratings to guide beneficiaries to higher-rated plans first. Medicare beneficiaries use the Star Ratings to assess the quality of the MA plans; and agents and brokers use the Star Ratings in assisting beneficiaries in selecting a plan that fits their health care needs.

47. Star Ratings thus influence each plan's position in the marketplace, by affecting how prospective enrollees, and the agents and brokers who advise them, perceive the comparative quality of various plans. For instance, MA-only plans with a 5.0 Star Part C summary rating and Part D plans with a 5.0 Star overall rating are displayed with a high-performing icon, while a plan that had any combination of Part C or Part D summary ratings of 2.5 Stars or lower in the most recent three consecutive years is marked with a "low performance" icon. *See* 42 C.F.R. §§ 422.166(h), 423.186(h).

48. The Star Ratings system is intended to reflect each plan's ability to provide quality care and benefits to its enrollees. It also affects the compensation that MAOs receive for the plans they sponsor. Moreover, regulations permit beneficiaries to change plans any time during the year, but only if the plan into which they move is a 5.0 Star plan. *Id.* § 422.62(b)(15). Star Ratings also drive whether a low-performing MA plan remains eligible to continue to participate in the program. *See Id.* §§ 422.502(b), 423.503(b).

D. Converting measure scores into Star Ratings: non-CAHPS measures

49. CMS uses two methodologies to convert measure scores into measure-level Star Ratings, depending on whether the underlying measures are drawn from the CAHPS survey (also called CAHPS measures) or non-CAHPS measures. For CAHPS measures, a method based on scores' relative percentile distribution and significance testing methodology is applied; for non-CAHPS measures, a clustering methodology is used. 42 C.F.R. §§ 422.166(a)(2), (3); 423.186(a)(2), (3).

50. The clustering methodology that CMS has used since the 2024 Star Ratings to convert measures scores to measure-level Star Ratings for non-CAHPS measures begins by removing "Tukey outer-fence outliers" from the set of scores for a given measure. *Id.* §§ 422.166(a)(2)(i), 423.186(a)(2)(i). This is a statistical technique used to "identify and remove extreme outliers in a dataset." *Elevance*, 2024 WL 2880415, at *5.

51. Once outlier deletion is done, CMS applies mean resampling with hierarchical clustering to sort the measure scores into groups. 42 C.F.R. §§ 422.166(a)(2)(i), 423.186(a)(2)(i). The "dividing lines" between the groups establish the Star Rating "cut points," which are the numerical values at which "a score results in a higher or lower star rating." *See SCAN*, 2024 WL 2815789, at *2; *Elevance*, 2024 WL 2880415, at *4. Clustering is a statistical technique used to partition a dataset into distinct groups, such that the observations within a group are as similar as possible to each other, and as dissimilar as possible to observations in any other group. *See* 42 C.F.R. § 422.162(a); 83 Fed. Reg. at 16525.

52. As a final step, CMS applies a "guardrail" to the cut points, capping any change in value from the previous year's cut points at five percent. 42 C.F.R. §§ 422.166(a)(2)(i), 423.186(a)(2)(i).

53. A plan's measure-level Star Rating for a given measure depends on where the plan's numerical measure score falls relative to the guardrail-capped cut points. Because

the measure-level Star Ratings are based on "whole star increments" (*id.* §§ 422.166(a)(4), 423.186(a)(4)), a change of even one percent in the cut points can lead to a whole-star drop in a plan's Star Rating.

E. Converting measure scores into Star Ratings: CAHPS measures

54. The relative distribution and significance testing methodology applied to convert CAHPS measure scores into Star Ratings begins by "case-mix adjust[ing]" the scores to "take into account differences in the characteristics of enrollees across contracts that may potentially impact survey responses." *2024 Technical Notes*, at 158.

55. Contracts are then classified into "base groups" by reference to "percentile cut points defined by the current-year distribution of case-mix adjusted contract means." 83 Fed. Reg. at 16568. These percentile cut points are set at the 15th, 30th, 60th, and 80th percentiles. *See* 42 C.F.R. §§ 422.166(a)(3), 423.186(a)(3). Each base group "includes those contracts whose rounded mean score is at or above the lower limit and below the upper limit" of these percentile cut points. 83 Fed. Reg. at 16568.

56. The Star Rating awarded to a contract for a given CAHPS measure depends on the contract's base group assignment and its interaction with the following factors: the statistical significance and direction of the difference between the contract mean and the national mean, the statistical reliability of the contract's measure score, and the standard error of the mean score. *Id.*; 42 C.F.R. §§ 422.166(a)(3), 423.186(a)(3). All statistical tests are computed using unrounded scores. 83 Fed. Reg. at 16568.

57. Reliability is a statistical property defined in this context as "the fraction of the variation among the observed measure values that is due to real differences in quality ('signal') rather than random variation ('noise')." 42 C.F.R. § 422.162(a). It is reflected in a scale from 0 (where all differences in measure scores are due to measurement error) to 1 (where all differences in measure scores are due to real differences in quality and perfor-

mance). *Id.* Where the reliability of a CAHPS measure is less than 0.60, it is designated a "very low reliability" measure (*2024 Technical Notes*, at 194), and "no measure Star Rating is produced." 42 C.F.R. § 422.166(a); *see 2024 Technical Notes*, at 160.

58. Because numerical measure scores are converted to measure-level Star Ratings with "whole star increments" (42 C.F.R. §§ 422.166(a)(4), 423.186(a)(4)), even minute fluctuations in the percentile cut points can change a contract's base group assignment and lead to a whole-star change in the contract's Star Rating.

F. The Foreign Language Interpreter and TTY Availability measure

59. Each MAO must be able to provide specific information on a timely basis to current and prospective enrollees upon request, including with a toll-free customer service call center. 42 C.F.R.§ 422.111(h). By regulation, call centers must limit average hold times to no more than two minutes, answer 80 percent of incoming calls within 30 seconds, and limit the disconnect rate of all incoming calls to no more than five percent, among other requirements. *Id.*

60. Among the measures underlying both Part C and Part D Star Ratings are the Foreign Language Interpreter and TTY Availability measures, which gauge the availability of teletypewriter (TTY) services and foreign-language interpretation to prospective enrollees who call a plan's customer service phone line speaking a language other than English. The measures are labeled C30 (underlying Part C ratings) and D01 (underlying Part D ratings). *See 2024 Technical Notes*, at 83, 85.

61. These measures monitor a plan's compliance with a CMS regulation that requires MAOs to maintain a customer service call center that makes foreign language interpreters available, at no cost to the caller, for "80 percent of incoming calls requiring an interpreter within 8 minutes of reaching the customer service representative." 42 C.F.R. §§ 422.111(h)(iii), 423.128(d)(1)(iii). The regulation does not specify that the eight-minute

requirement must be met in a single call and is silent on the permissibility of call backs in the event of a disconnected call.

62. CMS conducts the "Accuracy & Accessibility Study" to evaluate plan performance on these measures. The study uses surveyors to place anonymous test calls to plans' customer-service call centers. *See* CMS, *Memo: 2024 Part C and Part D Call Center Monitoring—Timeliness and Accuracy & Accessibility Studies* 1-2 (2024).

63. The metric used to assess foreign language interpreter availability is the number of *completed* interpreter contacts, divided by the number of attempted contacts during *connected* calls. *Id.* at 2. In line with sections 422.111(h)(1)(iii) and 423.128(d)(1)(iii), a call is considered connected when the "call surveyor" or test caller "reaches" a customer service representative. *Id.* A contact with an interpreter is considered "completed" when the caller "establish[es] contact with an interpreter and confirm[s] that the customer service representative can answer questions" about the plan's Medicare Part C or Part D benefits "within eight minutes." *2024 Technical Notes*, at 83, 85.

64. CMS has a practice of "invalidating" certain calls—that is, excluding them altogether from this ratio of completed to attempted contacts. CMS's practice is to invalidate calls when, among other circumstances, there is no evidence that the plan was at fault for a call that was not successfully connected or completed. CMS's study thus places a call into one of three categories: (1) successfully completed; (2) not successfully completed (after having been connected); or (3) invalidated (*i.e.*, excluded from the study and not considered for purposes of Star Ratings). Invalidating calls is not unusual during the plan preview process. As a result of Humana's dialog with CMS during the latest plan preview period, for instance, CMS invalidated four calls.

65. CMS test callers must follow predefined procedures before they may conclude that a call center has "completed" a call for assessment. First, the CMS test caller must

dial the plan number. Second, the test caller must connect with the plan's customer service representative. Third, the CMS test caller must ask an introductory question, to which the customer service representative must answer affirmatively.

66. In testing interpreter availability, the CMS test caller will place a call to the plan's customer service call center in a foreign language and wait for the customer service representative to bring an interpreter to the phone to assist the representative in answering the introductory question. CMS allows for an eight-minute window for the customer service representative to connect to an interpreter and answer the introductory question and questions about plan benefits. See *Timeliness and Accuracy & Accessibility Studies*, at 2. These prerequisites ensure that the call center is evaluated according to its own actions or inactions and not assigned responsibility for problems outside its control.

67. To receive 5.0 Stars on the call center measure in the 2025 Star Ratings, CMS required 100% of non-invalidated foreign language calls to be scored as successful. Given the demand for perfection to receive 5.0 Stars on the call center measure, CMS's decisions regarding whether and how to score just a single call included in the study can have an enormous impact on a plan's overall Star Rating, and consequently an outsized impact on a plan's ability to offer competitive benefits and premiums for enrollees.

G. Plan sponsor participation in Star Ratings (plan preview periods)

68. Given the importance of Star Ratings to the MA program, and the sensitivity of the system to erroneous or unreliable data, CMS's regulations establish an administrative process through which MAOs and other plan sponsors can review and comment on, and challenge the adequacy of, the agency's preliminary calculations. The regulations call this administrative process the "plan preview" periods: "CMS will have plan preview periods before each Star Ratings release during which MA organizations can preview their Star Ratings data in HPMS prior to display on the Medicare Plan Finder." 42 C.F.R. § 422.166-

(h)(2); *see also id.* § 423.186(h)(2). HPMS is CMS's Health Plan Management System, a website used to facilitate communications between CMS and MAOs. *See* https://-hpms.cms.gov/app/ng/home/.

69. The plan preview process is the only administrative process available to a plan permitting it to comment on and participate in the Star Ratings process before Star Ratings are finalized and published by CMS.

70. CMS holds two preview periods. During the first plan preview, CMS "expect[s]" to "closely review the methodology and their posted numeric data for each measure." 83 Fed. Reg. at 16588. During the second plan preview, CMS will post to the System "any revisions made as a result of the first plan preview," as well as the "preliminary Star Ratings for each measure, domain, summary score, and overall score." *Id.* CMS again "expect[s]" plan sponsors to "closely review the methodology and their posted data for each measure, as well as their preliminary Star Rating assignments." *Id.*

71. A core purpose of the plan preview process is data validation. The two plan previews allow "sponsors to review and raise any questions about their own plan's data prior to the public release of data for all plans," so that if there are any errors, "necessary corrections" can be made prior to the Star Ratings being announced to the public. *Id.*

72. The plan preview process reflects CMS's position that, for the Star Ratings to be a "true reflection of the quality, performance and experience of the beneficiaries enrolled in MA and Part D contracts," the data and analysis underlying measure scores and measure-level Star Ratings must be "complete, accurate, and unbiased." *Id.* at 16567. Because of the importance of data accuracy, "[d]ata validation is a shared responsibility among CMS, CMS data providers, contractors, and Part C and D sponsors." *Id.* at 16562.

73. CMS imposes harsh penalties on MAOs for submitting inaccurate data. For example, CMS will "reduce a contract's measure rating when CMS determines that" the

data reported to it "are inaccurate, incomplete, or biased," which can result from "mishandling" and "inappropriate processing" of data. 42 C.F.R. §§ 422.164(g)(1), 423.184(g)(1). For measures based on data that an MAO must submit to CMS, the rating will be reduced to 1 Star when a contract "was not compliant with CMS data validation standards." *Id.* §§ 422.164(g)(1)(ii), 423.184(g)(1)(ii).

FACTUAL ALLEGATIONS

A. CMS's refusal to disclose data in the plan preview process

74. Although CMS has expressed a firm commitment to ensuring the reliability and accuracy of the data it uses to calculate the Star Ratings, Humana's experience paints a different picture. Humana has identified Star Rating calculation errors during the plan preview periods on multiple occasions. Humana's comments, made to ensure accuracy, have rarely been met with agency resistance in the past.

75. In 2015, for example, Humana observed a large percentage of CAHPS respondents indicating their healthcare services were not provided by Humana. When Humana presented its concern to CMS, the agency responded with what appeared to be cut-and-paste email language insisting that the samples had been "verified" by the agency for all contracts and did not include enrollees that had been in the plan less than six months. That was untrue. Humana was able to determine from vendor data that the sample included many new enrollees who did not meet the six-month continuous enrollment requirement. CMS later conducted an investigation—the actual verification it claimed to have undertaken earlier but did not—acknowledged its error, and had to recalculate CAHPS measure results for the entire industry.

76. The systemic error uncovered in 2015 was not an entirely isolated incident. During the plan preview period in 2021, for instance, Humana's internal data regarding enrollee appeals did not match the measure rates provided by CMS. Humana discovered

that the agency had omitted an entire day—December 31, 2020—from its data. The agency later confirmed and corrected the mistake. And earlier this year, Humana was unable to validate rates for the "Members Choosing to Leave the Plan" measure for its largest contract, H5216. When CMS later disclosed its data, Humana determined that the agency had erroneously identified enrollees moving from two smaller contracts that had been merged into H5216 as "choosing to leave the plan" when they had not. Humana raised the issue with CMS, which confirmed its mistake.

77. To be fair, due to the complexity and volume of data required to perform Star Ratings calculations, errors in CMS's data and calculations are inevitable. But especially given the importance of the Ratings to the MA program, CMS has an obligation to ensure that every detail is *truly* checked and double-checked. That is why it is so important for the agency to make available all data necessary for plan sponsors' validation work, as required by 42 C.F.R. § 422.166(h)(2)—and, indeed, why it is arbitrary and capricious for the agency to do otherwise.

78. In a break from past practice, CMS has resisted many of Humana's requests and comments during the plan preview periods for the 2025 Star Ratings. This appears to be part of an emerging trend, as suggested by the filing of five other lawsuits in 2024 alone, all challenging the agency's Star Ratings methodology. Two of those lawsuits, brought by SCAN Health Plan and Elevance Health, already have resulted in orders from other courts against the agency, requiring a retraction of each company's 2024 Star Ratings. Lawsuits of this kind, nearly unheard of before this year, have become a regrettable necessity.

79. In this case, CMS arbitrarily declined to make critical score-validation data available to Humana and refused to treat plans consistently with respect to other measures. During the second plan preview period in late September 2024, Humana requested numer-

ous data essential to validate CMS's preliminary calculations. By email dated September 17, 2024, Humana explained that it had

> sent several emails over the course of the Star Ratings plan preview ("PP") 2 period to [CMS], asking for additional information related to Humana's 2025 Star Rating performance and threshold calculations. In many cases, CMS has not provided the requested data, making it impossible for Humana to fully assess suspected errors in CMS's Star Rating calculations for certain measures. CMS also has not articulated any reasonable bases for failing to provide the necessary data. With PP2 closing later today, Humana has significant concerns about the reliability of the Star Rating data and calculations underlying its ratings and objects to the finalization of its Star Ratings without CMS providing this data.

The email went on to explain that "Humana has upheld its obligation to quickly and rigorously evaluate the data provided by CMS and has in fact identified numerous suspected errors," but that Humana's data- and calculation-validation efforts were "met with resistance from the agency," which "failed to provide the necessary data for Humana to complete its review." It continued:

> Our concern that additional errors exist is well-founded. For many measures that are normally very stable, we have noticed threshold movement inconsistent with average industry performance and historical trends, including but not limited to: the 2-star threshold for the MPF Price Accuracy measure (D07) in Star Rating year 2025 being equal to the 5-star threshold in Star Rating year 2023; the 4 and 5-star thresholds for the Medication Reconciliation Post-discharge measure (C14) increasing by 5 percentage points after barely moving at all in the past three rating years; and all threshold levels for the Diabetes Care—Blood Sugar Controlled measure (C10) increasing at much higher rates than industry performance over the past two rating years. We have also noticed an unusually high number of 5-star thresholds at 99 and 100 percent. We are concerned that calculation errors could be driving these irregularities. . . . Moreover, in developing CAHPS thresholds, we believe CMS failed to exclude both statistically unreliable measure scores and measure scores from disaster-impacted contracts.

To this email, Humana attached prior communications in which it had asked for specific data necessary for its own validation of CMS's work concerning these issues, with explanations of the need.

80. Prior requests for data that had gone unanswered included requests for:

- the de-identified industry contract rates for all contracts with more than 10 valid responses, for all 2024 and 2025 CAHPS measures except the Rating of Health Plan measure
- the de-identified industry data and unrounded cut point rates for the following measures: Medication Reconciliation Post Discharge; Diabetes Care—Blood Sugar Controlled; Special Needs Plan (SNP) Care; Medicare Plan Finder: Price Accuracy
- the de-identified and unrounded industry rates to validate the cut-point calculations for measures derived from the Health Outcome Survey
- the enrollee-level data file of survey participants, which would allow Humana to validate that enrollees included in the CAHPS results meet all eligibility requirements
- the de-identified industry data and unrounded cut-point rates for the following measures: Part C Interpreter and TTY Functionality; Part D Interpreter and TTY Functionality.

81. CMS did not disclose these data. It instead provided a zip file of a means- and test-report data and a case-mix report for each Humana contract. CMS also provided a file of de-identified scores for one measure—the Rating of Health Plan measure—for all contracts. But it did not provide any of the data above described. Without this information, Humana was unable to validate CMS's cut-point calculations, its CAHPS survey data, or more generally its measure-level Star Ratings.

82. CMS's explanation for declining to make the data available was as follows:

- Concerning cut-points, it said simply that "changes in the distribution of scores across contracts each year impact the cut points," and "[t]his does not indicate data inaccuracies."
- Concerning CAHPS measures, the agency stated: "We do not believe that additional member-level details for all contracts are needed to successfully validate scores" and that "[i]t is CMS policy that plans should not receive any identifiable MA & PDP CAHPS survey data."
- Concerning Health Outcome Survey measures, CMS stated that it could not provide the requested data until follow-up data collection was complete because to do so would compromise "the integrity of the sample."

B. Concerning patterns in CMS's data and calculation results

83. CMS's refusal to disclose the data needed for Humana to validate the cut points is especially concerning given the few disclosures CMS *has* made. For example, during the second plan preview period, CMS assured Humana that, as required by regulation, CAHPS measure scores that are calculated with "very low reliability" for a given measure are not assigned a measure-level Star Rating and are not included when determining a *contract's* summary and overall ratings. But CMS then indicated to Humana that CAHPS measure scores for plans with "very low reliability" data *are* included in both the calculation of base group cut points and the national mean used in significance testing. In other words, CMS is using unreliable data and not comparing apples to apples. The agency has declined to explain why "very low reliability" is properly excluded from single-contract calculations but *not* national cut points and significance testing.

84. More recently, after the public 2025 Star Ratings data were officially released, Humana attempted to replicate the cut point calculations combining the limited data shared in the plan preview periods with the data publicly available as of October 10, 2024. Humana observed that 60% of its replicated calculated cut points did not match the CMS published cut points, with some varying by as much as 14 percentage points, an enormous difference.

85. Using this same data and the mean resampling methodology that CMS reports using, Humana conducted ten random runs of cut point calculations. These produced results varying by as much as 6 percentage points for a single measure-level star cut point and the averages varying from the published cut points—again, a massive difference.

86. Applying the Tukey outlier deletion methodology, Humana also was unable to validate the upper and lower bounds for the data used to calculate cut points. Its results

matched CMS's only about half of the time, and some showed wide swings from the lower outlier in particular: up to 8 percentage points different from CMS's published results.

87. In undertaking these validation efforts, Humana further determined that the data shared for four sample measures during the plan preview periods does not match the results that were published. For example, the sample data shared during Plan Preview 2 for the Breast Cancer Screening measure showed 541 records, whereas the data published in the industry files only shows only 500 records.

88. Each one of these discrepancies alone would raise serious concern that the agencies methodologies and calculations were applied incorrectly. Collectively, they leave little doubt that the cut points are inaccurate.

89. CMS's approach to third-party validation of the data, methodology, and calculations for the 2025 Star Ratings has violated the plain terms and clear policy of its own regulations, including 42 C.F.R. §§ 422.166(h)(2) and 423.186(h)(2), which require the agency to hold plan preview periods in which "MA organizations can preview their Star Ratings data"; and 42 C.F.R. §§ 422.164(g)(1) and 423.184(g)(1), which commit the agency—every bit as much as MAOs—to ensuring that there is no "mishandling of data, inappropriate processing, or implementation of incorrect practices that have an impact on the accuracy, impartiality, or completeness of the data used for" the Star Ratings, and more generally that "measure data are [not] inaccurate, incomplete, or biased."

C. Disconnected calls under the Accuracy & Accessibility Study

90. During the CY 2025 Accuracy & Accessibility Study, CMS identified two relevant calls (case IDs D1100955 and D0900533) placed to Humana customer service representatives as "incomplete" because the calls disconnected. The two calls disconnected due to third-party internet connection interruptions while Humana's service represen-

tatives were connecting with an interpreter to join the call. Together, these calls reduced the overall Star Ratings for many of Humana's largest contracts.

91. CMS monitors call disconnects under the separate Timeliness Study element of the Call Center Monitoring Program.

92. The two calls disconnected while Humana's customer service representatives were actively connecting with an interpreter to join the call. In the event of a dropped call, Humana's standard protocol is for the customer services representative to call the prospective enrollee back. But CMS call surveyors do not accept attempts to call back after initial calls are disconnected. Because these two calls dropped, the call surveyors designated them "incomplete."

93. In response to inquiries from Humana during the second plan preview period for the 2025 Star Ratings, CMS indicated that—as a policy—it does not allow callbacks from plans in the context of the Accuracy & Accessibility Study, requiring the call surveyors to receive responses to their questions in the test language within a single call. When Humana explained its standard protocol in response to dropped calls, CMS responded this way in an email sent on September 16, 2024:

> Your plan is disputing these calls as your procedure is to obtain the phone number of the prospective member and then to call them back if there is a disconnection. However, CMS does not allow callbacks from the plan as all questions should be answered in a single call. . . . [It also] will not revise results based on challenges to the methodology, which has been applied to all subjects of the study.

94. CMS has never explained why "all questions should be answered in a single call," or why this single-call criterion should be given controlling significance. The agency's regulations require only that a plan make foreign language interpreters available to assist non-English speaking (and limited English proficient) Medicare beneficiaries within eight minutes of an initial connection. 42 C.F.R. §§ 422.111(h)(1)(iii), 423.128-

(d)(1)(iii). The regulations do not establish a single-call requirement, which was not adopted following notice-and-comment rulemaking.

95. If the CMS call surveyors were permitted to accept callbacks, Humana's customer service representatives would have called back and been able to answer the callers' questions with the help of an interpreter within eight minutes of originally connecting with the callers, as required by 42 C.F.R. §§ 422.111(h)(1)(iii) and 423.128(d)(1)(iii). If CMS had not imposed additional requirements that do not appear in the Code of Federal Regulations and were not subjected to notice-and-comment, the calls at issue would have been designated "complete" for purposes of the Accuracy & Accessibility Study.

96. CMS's arbitrary policy of not permitting callbacks adds unlawfully to the requirements for foreign language calls appearing in sections 422.111(h)(1)(iii) and 423.128(d)(1)(iii), which is connection with an interpreter within eight minutes. Moreover, the policy does not improve the "quality" of a plan or its services, and it leads to double-counting of technical call drops in the context of foreign-language calls—such a dropped call counts against a plan sponsor in both the Timeliness Study and the Accuracy & Accessibility Study.

97. CMS also has treated calls like D0900533 and D1100955 differently across plan sponsors, despite analytically indistinguishable facts. For instance, in 2023, the health insurer Elevance Health and its affiliates, all using the same call center, challenged a CMS determination evaluating them under the same criterion at issue here, measure D01, the "Call Center—Foreign Language Interpreter and TTY" metric. CMS concluded that their call center had missed a single call, and on that basis awarded a lower overall Star Rating for Elevance's plans for 2024.

98. Following litigation in the U.S. District Court for the District of Columbia, CMS resolved the dispute in favor of Elevance because there was no evidence that Elevance

was responsible for the call being dropped. Elevance described the final CMS determination as follows: "Based on the evidence presented by Elevance and CMS, the CMS Reconsideration Official found that there was no evidence the call at issue failed due to actions by Elevance and should not have counted against Elevance." The agency thus awarded the plan a higher Star Rating.

99. Just as in the Elevance case, there is no evidence that calls D0900533 or D1100955 failed due to actions by Humana, and they accordingly should not have counted against Humana. CMS did not acknowledge or attempt to explain this disparate treatment under the Accuracy & Accessibility Study.

D. Call in which no question was asked

100. CMS identified a third call (case ID C0701002) as "incomplete" during the CY 2025 Accuracy & Accessibility Study, which again adversely impacted the Star Ratings calculations for Humana's largest contracts. During this test call to Humana's call center, neither the Humana representative nor the CMS call surveyor spoke; each was silent throughout the entire duration of the call. After an extended period of silence with no communication, the call was disconnected.

101. CMS requires plan sponsors to make foreign language interpreters available to would-be enrollees "within 8 minutes of *reaching* the customer service representative." 42 C.F.R. §§ 422.111(h)(iii), 423.128(d)(1)(iii) (emphasis added). CMS guidance regarding the Accuracy & Accessibility Study confirms the agency's view that a call is "connected" only when the call surveyor "reaches" a customer service representative. AR85.

102. In the context of a telephone call, the word "reach" means *communicate with*. *See New Oxford American Dictionary* 1415 (2001) (defining "reach" as to "communicate with (someone) by telephone or other means"); *Webster's Third New International Dictionary* 1888 (1986) (defining "reach" as to "communicate with . . . by phone"). When a call

surveyor does not successfully communicate even a single word with a customer service representative, the call surveyor has not "reached" the representative.

103. Even when a call technically is connected, CMS guidelines provide that the test caller must then pose an "introductory question" to "to ensure [that the caller is] speaking with a representative in the correct department." AR82. An example of an introductory question is, "Are you the right person to answer questions about [Plan name's] health benefits?" *Id.* If the answer is "yes," the call moves to the next phase, in which the caller indicates a need to speak with a foreign-language speaking representative. *Id.*

104. Because the CMS call surveyor in call ID C0701002 did not communicate with the Humana customer service representative, the CMS caller did not "reach" the Humana customer service representative.

105. Even if a connection technically was made, the CMS call surveyor failed to ask an introductory question, that agency's own guidelines thus were not followed, and CMS should have invalidated the call for purposes of the Accuracy & Accessibility Study.

E. Unlawful sub-delegation of agency authority

106. As since revealed by CMS's incomplete administrative record in this case, the agency has also improperly delegated decisionmaking authority to a private subcontractor. CMS simply forwarded Humana's communications concerning the challenged calls to the subcontractor, who proposed denying Humana's objection. CMS then relayed the contractor's decision to Humana, passing it off as its own.

107. For example, when CMS forwarded Humana's objection concerning callbacks and case IDs D1100955 and D0900533, the contractor replied on September 16, 2024, at 9:44am: "[W]e do not allow callbacks from the plan as all questions should be answered in a single call." CMS then replied to Humana approximately four hours later: "CMS does not allow callbacks from the plan as all questions should be answered in a single call."

108. In another example, when CMS forwarded Humana's objection concerning case ID C0701002, the subcontractor stated that "before the CMS caller could make contact with a CSR the call disconnected." Because "the disconnect was not initiated by the CMS caller," the contractor "[r]ecommend[ed] keeping outcome as is." One hour later, CMS responded without explanation, "I agree with keeping the call as is." CMS then responded to Humana approximately 90 minutes later, stating "Call C0701002 will remain as is" because "before the CMS caller could make contact with a CSR the call disconnected," and "the call log shows that the disconnect was not initiated by the CMS caller."

F. Final agency action and harm to plaintiffs

109. CMS issued the final 2025 Star Ratings on October 10, 2024. The determination of the 2025 Star Ratings is final, not tentative; CMS does not require or provide for appeals or other internal review of Star Ratings methodology. The plan preview periods are the last opportunity that a plan sponsor may or must use to administratively challenge an adverse change in a contract's Star Rating from one year to another resulting from the use of erroneous cut points or otherwise unlawful calculation methodologies. *See* 42 C.F.R. § 422.260(c)(3)(ii) (providing for administrative review of QBP determinations but excluding challenges to the "methodology for calculating the star ratings" and the "cut-off points for determining measure thresholds").

110. A final Star Rating determines legal rights and obligations, and legal consequences flow from them. For example, CMS may terminate a plan's MA contract that has failed to achieve a Part C summary rating of at least three stars for three consecutive contract years. 42 C.F.R. § 422.510(a)(4)(xi). In addition, while plans are typically barred from allowing Medicare beneficiaries to switch to their plan until the annual enrollment period, regulations permit such a switch at any time during the year if the plan into which a beneficiary moves has a 5.0 Star Rating. *Id.* § 422.62(b)(15).

111. Issuance of Star Ratings also immediately injures adversely affected plan sponsors like Humana. When the MA annual enrollment period begins on October 15, 2024 (following publication of all Star Ratings in Medicare Plan Finder) the Star Ratings impact each plan's reputation. Humana may suffer reputational injury given its previous track record of earning industry-leading Star Ratings.

112. Under CMS's 2025 Star Ratings calculations, more than one dozen Humana MA contracts were near the cut point from 3.5 Stars to 4.0 Stars or the cut point from 4.0 Stars to 4.5 Stars. Upon information and belief, these contracts would have received higher Star Ratings had the cut points been correctly calculated. However, Humana was not permitted to validate, and therefore could not fully challenge, the calculations because CMS denied access to the data necessary for doing so.

113. Upon information and belief, Humana's largest contracts would have received higher Star Ratings had CMS not imposed an irrational, unexplained, and unlawful "single call" requirement for compliance with 42 C.F.R. §§ 422.111(h)(iii) and 423.128(d)(1)(iii).

114. ABC's members also are injured directly by the 2025 Star Ratings. Using unvalidated data and methodology to derive the highest cut points in program history, CMS's 2025 Star Ratings have reduced the number of plans with 4.0+ Stars to the lowest level in recent memory. For the past five years, for example, the proportion of MA enrollees in 4.0+ plans ranged between 78% to 91%, reflecting the high quality of MA plans across the industry. The proportion of MA enrollees who will be in 4.0+ plans under the 2025 Ratings is now just 64%, but with no indication that MA plan quality has meaningfully declined. As a result, countless MA enrollees—many, members of ABC—will now find themselves in plans with less generous QBPs and thus with less ability to include supplemental benefits, like vision, hearing, and dental coverage.

115. Agents and brokers also are injured. Again, a key aim of the Star Ratings system is to offer Medicare beneficiaries and their agents and brokers "comparative information on plan quality and performance" to allow them to make "knowledgeable enrollment and coverage decisions." 42 C.F.R. §§ 422.160(b)(1), 423.180(b)(1). Agents, brokers, and beneficiaries' ability to use the Star Ratings program to inform their enrollment and coverage decisions depends on the ratings being a "true reflection" of plan quality and performance. 83 Fed. Reg. at 16520. When Star Ratings are not grounded in validated data or sound methodologies, they cannot be taken to reflect accurately the quality and performance of plans. Agents and brokers working to make the best and most accurate recommendations for their clients are therefore now having to do additional background research on the plans they recommend, to confirm whether or not plans whose Stars Ratings have been lowered for 2025 actually remain high quality plans.

116. More broadly, agents, brokers, and beneficiaries are harmed when CMS does not use validated data and methodologies to calculate Star Ratings, undermining the integrity and reliability of the Star Ratings program.

CLAIMS FOR RELIEF

Count I
Refusal to Disclose Information Needed for Validation

117. Plaintiffs reallege the allegations set forth in all prior paragraphs of this complaint as though fully set forth herein.

118. By regulation, CMS has committed to making data available as necessary for MAOs to validate CMS's data and all relevant calculation for its Star Rating determinations. 42 C.F.R. § 422.166(h)(2). Here, CMS refused to share all of the data required for plan sponsors to validate the agency's Star Ratings calculations, arbitrarily deciding to disclose only a limited range of agency-selected information.

119. It is arbitrary and capricious for CMS to make only some, but not all, relevant data available. Agencies bear a general obligation to disclose all data essential to their decision-making. *Chemical Manufacturers Association v. EPA*, 870 F.2d 177, 200 (5th Cir. 1989) ("[F]airness requires that the agency afford interested parties an opportunity to challenge the underlying factual data relied on by the agency."); *Texas v. EPA*, 389 F. Supp. 3d 497, 505 (S.D. Tex. 2019) (quoting *American Radio Relay League, Inc. v. FCC*, 524 F.3d 227, 237 (D.C. Cir. 2008)).

120. "[I]t is elementary that an agency must adhere to its own rules and regulations. *Ad hoc* departures from those rules . . . cannot be sanctioned." *Texas v. EPA*, 829 F.3d 405, 430 (5th Cir. 2016) (quoting *Reuters Ltd. v. FCC*, 781 F.2d 946, 950 (D.C. Cir. 1986)). It is furthermore arbitrary and capricious for an agency to pursue a policy (non-disclosure of data) that conflicts with an applicable regulation (requiring disclosure). *State v. EPA*, 91 F.4th 280, 291 (5th Cir. 2024) (citing *Gulf States Manufacturers, Inc. v. NLRB*, 579 F.2d 1298, 1308 (5th Cir. 1978)).

121. CMS's refusal to disclose the data that Humana requested is at odds with its policy of seeking the participation of plan sponsors in the validation of its empirical analyses. *See* 42 C.F.R. §§ 422.166(h)(2); 423.186(h)(2). It also conflicts with CMS's strict data integrity standards for plans. *See id.* §§ 422.164(g)(1), 423.184(g)(1).

122. CMS's policy of denying access to information needed to validate its cut point calculations frustrates the overall congressional policy of establishing a data-driven quality rating system based on accurate data validated by plan sponsors.

123. CMS did not provide meaningful responses to Humana's requests for data. Using apparently cut-and-paste language, it asserted simply that its cut point calculations were accurate, and that further data was unnecessary to perform data validation tasks.

124. The only further explanation given by CMS concerned the Health Outcome Survey measures. On that front, the agency asserted that it could not provide the requested data until follow-up data collection was complete because to do so would compromise "the integrity of the sample." But this same data was used by CMS to calculate the Star Ratings. CMS has not provided a rational explanation why providing the same data to Humana for data validation would impermissibly compromise the integrity of the sample.

125. The substantial swings in the 2025 cut points relative to recent prior years introduces a plausible concern that the 2025 cut points are inaccurate. This may be because the data underlying the agency's calculations is flawed, or because its analysis is flawed. Either way, there are no independent, objective indications that MA plan quality has diminished in recent years as suggested by the decline in overall star ratings across the program, and the cut points are therefore arbitrary and capricious.

126. Accordingly, Humana's 2025 Star Ratings should be set aside. CMS must issue new 2025 Star Ratings according to fully validated and confirmed data and statistical analyses, in accordance with the APA and the agency's regulations.

Count II
Sub-delegation to Nongovernmental Third Party

127. Plaintiffs reallege the allegations set forth in all prior paragraphs of this complaint as though fully set forth herein.

128. "An agency abdicates its role as a rational decision-maker, and impermissibly subdelegates, if it does not exercise its own judgment, and instead cedes near-total deference to private parties'" judgments. *Louisiana Public Service Commission v. FERC*, 761 F.3d 540, 551 (5th Cir. 2014) (cleaned up). "Agencies must *actually exercise* their authority rather than reflexively rubber stamping [decisions] prepared by others." *Consumers' Research Cause Based Commerce v. FCC*, 109 F.4th 743,770 (5th Cir. 2024) (cleaned up).

129. To delegate agency decisionmaking to a nongovernmental third party is per se arbitrary and capricious and a violation of statutory duties.

130. CMS improperly subdelegated discretionary agency authority over Humana's objections during the plan preview process and more generally over the regulatory requirements for the Accuracy & Accessibility Study and call center measures. This was not in keeping with statutory requirements and was inherently arbitrary and capricious.

131. Accordingly, Humana's 2025 Star Ratings should be set aside. CMS must issue new 2025 Star Ratings that are calculated without reliance on a private contractor's recommendations.

Count III
Disparate Treatment of Similar Call Disconnections

132. Plaintiffs reallege the allegations set forth in all prior paragraphs of this complaint as though fully set forth herein.

133. An agency must regulate in consistent ways that do not vary arbitrarily from on regulated party to another. If MA plan sponsors' Star Ratings are predicated on different standards of decision, the Star Ratings themselves lose their reliability as a measure permitting plan-to-plan comparisons.

134. In prior cases, CMS has invalidated an "incomplete" call when there is no evidence the call at issue failed due to actions by the MA plan. Here, there is no evidence that calls D0900533 or D1100955 failed due to actions by Humana, and they accordingly should not have counted against Humana under the standard that CMS has applied in other cases. Yet CMS counted the calls against Humana anyway. It is arbitrary and capricious for CMS to treat similarly situated plan sponsors differently without an evidentiary basis or rational explanation.

135. The arbitrarily disparate treatment of similarly situated plans undermines the reliability of the Star Ratings system.

136. Accordingly, Humana's 2025 Star Ratings for all contracts impacted by calls D0900533 or D1100955 should be set aside. The Court should remand the matter to the agency with directions not to include calls D0900533 or D1100955 as incomplete calls under the CY 2025 Accuracy & Accessibility Study.

Count IV
No-Callback Policy

137. Plaintiffs reallege the allegations set forth in all prior paragraphs of this complaint as though fully set forth herein.

138. CMS's Accuracy & Accessibility Study is conducted to monitor compliance with CMS regulations that require plan sponsors to make foreign language interpreters available to non-English speaking and limited English proficient prospective enrollees "within 8 minutes of reaching the customer service representative" by telephone. 42 C.F.R. §§ 422.111(h)(iii), 423.128(d)(1)(iii). The regulation does not specify that the eight-minute requirement must be met in a single telephone call and is silent on the permissibility of callbacks.

139. CMS guidance has consistently indicated that a call with an interpreter is defined as "completed" when the caller is able to receive responses to their questions about plan benefits "within eight minutes." *2024 Technical Notes*, at 83, 85; *Timeliness and Accuracy & Accessibility Studies*, at 2. The guidance is also silent on callbacks.

140. CMS has adopted a policy without notice-and-comment rulemaking by which a call will be deemed "incomplete" if it is disconnected, even if the call surveyor receives a callback and is able to receive responses to his or her questions with the help of an

interpreter less than eight minutes after initially reaching a customer service representative. CMS's policy is not to permit callbacks under any circumstance.

141. In the disconnected calls that Humana challenges in this suit, the customer service representative was in the process of connecting the call surveyor with an interpreter when the call dropped due to technical reasons outside of Humana's control. Had the customer service representative been permitted to call back after the call was disconnected, the call surveyors would have received responses to their questions within eight minutes of initial contact.

142. CMS's no-callback policy violates its regulations and guidance, which allow an eight-minute timeframe from the time a caller reaches a customer service representative for the caller to receive responses to questions about plan benefits, regardless of whether this happens within a single call. CMS's policy is thus arbitrary and capricious—as are, by extension, the CY 2025 Accuracy & Accessibility Study scores for Humana impacted by calls D0900533 or D1100955, and the ultimate 2025 Star Rating for those contracts. *See State v. EPA*, 91 F.4th at 291.

143. CMS's no-callback policy further violates 42 U.S.C. § 1395hh(a), which requires every "rule, requirement, or other statement of policy" that "establishes or changes a substantive legal standard" that determines either "payment for services" or "the eligibility of individuals, entities, or organizations to furnish or receive services or benefits" to be adopted by notice-and-comment rulemaking. The no-callback policy was not adopted in conformity with that requirement.

144. The no-callback policy also double counts disconnects. CMS's separate Timeliness Study assesses disconnect rates for calls to MAO customer service call centers. *See Timeliness and Accuracy & Accessibility Studies*, at 1. By also including disconnect rates within the Accuracy & Accessibility Study, CMS double counts technical call drop issues

related to foreign-language phone calls, artificially depressing plans' Star Rating scores. It is arbitrary and capricious for CMS to double count a call drop in this way.

145. Accordingly, Humana's 2025 Star Ratings for contracts impacted by calls D0900533 or D1100955 should be set aside. The Court should remand the matter to the agency with directions not to include calls D0900533 or D1100955 as incomplete calls under the CY 2025 Accuracy & Accessibility Study.

Count V
Study Call Where No Question Was Posed

146. Plaintiffs reallege the allegations set forth in all prior paragraphs of this complaint as though fully set forth herein.

147. CMS requires plan sponsors to make foreign language interpreters available to would-be enrollees "within 8 minutes of *reaching* the customer service representative." 42 C.F.R. §§ 422.111(h)(iii), 423.128(d)(1)(iii) (emphasis added). CMS guidance regarding the Accuracy & Accessibility Study confirms the agency's view that a call is "connected" only when the call surveyor "reaches" a customer service representative.

148. In the context of a telephone call, the word "reach" means *communicate with*. *See New Oxford American Dictionary*, at 1415; *Webster's Third*, at 1888. When a call surveyor does not successfully communicate with a customer service representative, the call surveyor has not "reached" the representative, and the call has not been "connected."

149. Even if the call is technically connected, CMS guidelines require the survey caller first to pose an introductory question of the customer service representative before proceeding with further steps to reach a foreign-language speaking representative. When no introductory question is posed, the surveyor cannot proceed with those further steps. Such a call therefore must be invalidated.

150. Under 42 C.F.R. §§ 422.111(h)(iii), 423.128(d)(1)(iii) and applicable CMS guidance, call C0701002 should be invalidated. CMS's decision (if it was the agency's decision) not to invalidate the call was arbitrary and capricious and contrary to law.

151. Accordingly, Humana's 2025 Star Ratings for contracts impacted by call C0701002 should be set aside. The Court should remand the matter to the agency with directions to invalidate call C0701002 and refrain from counting it as an incomplete call under the CY 2025 Accuracy & Accessibility Study.

PRAYER FOR RELIEF

WHEREFORE, plaintiffs ask the Court to enter judgment in their favor and:

(a.) Set aside and vacate Humana's 2025 Star Ratings and remand the matter to CMS for recalculation of Humana's 2025 Star Ratings without application of the unlawful practices and policies identified above;

(b.) Declare that CMS's policy refusing to disclose all relevant data and information necessary to permit MAOs to validate the data and calculations underlying CMS's Star Ratings is arbitrary, capricious, and unlawful;

(c.) Set aside and vacate Humana's 2025 Star Rating for all contracts adversely impacted by call IDs D0900533 or D1100955 on the ground that CMS's policy of refusing callbacks after dropped calls is unlawful, and remand to the matter to CMS;

(d.) Set aside and vacate Humana's 2025 Star Rating for all contracts adversely impacted by call ID C0701002 on the ground that a call surveyor does not reach a customer service representative if no communication takes place, and remand the matter to CMS;

(e.) Declare that CMS's policy of refusing callbacks after dropped calls and practice of refusing to invalidate calls lacking communication is unlawful for purposes of the Accuracy & Accessibility Study;

(f.) Award plaintiffs their costs and attorneys' fees as permitted by law;

(g.) Award plaintiffs such other and further relief as the Court may deem just and proper.

Dated: November 27, 2024

Respectfully submitted,

/s/ *Michael B. Kimberly*

Michael B. Kimberly* (D.C. No. 991549)
Kate McDonald* (D.C. No. 998233)
Caleb H. Yong* (D.C. No. 1780922)
Nicole E. Wittstein* (D.C. No. 1725217)
McDermott Will & Emery LLP
500 North Capitol Street NW
Washington, D.C. 20001
(202) 756-8901
mkimberly@mwe.com
kmcdonald@mwe.com
cyong@mwe.com
nwittstein@mwe.com

Richard Salgado (Texas No. 24060548)
McDermott Will & Emery LLP
2501 North Harwood Street, Suite 1900
Dallas, TX 75201-1664
(214) 210-2797
richard.salgado@mwe.com

* *pro hac vice*

Counsel for Plaintiffs

CERTIFICATE OF SERVICE

Undersigned counsel certifies that a true and correct copy of this document was served via CM/ECF on all counsel of record pursuant to the Federal Rules of Civil Procedure on November 27, 2024.

/s/ *Michael B. Kimberly*

Michael B. Kimberly